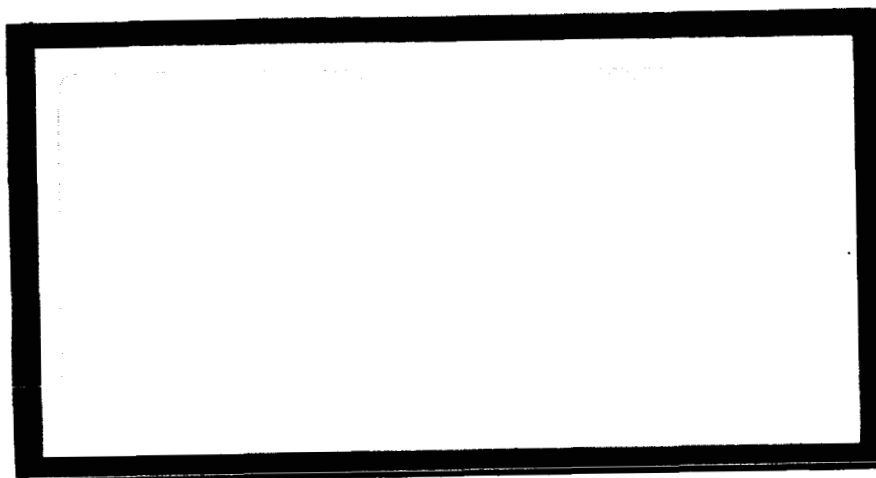


68p.

AD/COM
N64 12157
CODE-1
CR-55086



OTS PRICE

XEROX

\$

6.60 pd

MICROFILM

\$

2.24 mf.

ADCOM, INC.
238 Main Street,
Cambridge 42, Mass.
UN 8-7386

ADVANCED COMMUNICATION RESEARCH AND DEVELOPMENT

SELECTION OF SIGNALS
FOR
SPACE-VEHICLE COMMAND

→ (June 1, 1963 68p (0-19)

Elie J. Baghdady

prepared for

National Aeronautics and Space Administration
George C. Marshall Space Flight Center
Huntsville, Alabama

under

[2]

(NASA Contract No. NAS 8-2668)

(NASA CR-55086) OTS: \$6.60pr, \$2.24 mf

Submitted by

0079601

ADCOM, Inc.,
238 Main Street

Cambridge, ~~MA~~ Massachusetts

TABLE OF CONTENTS

| | <u>page</u> |
|---|-------------|
| I. Introduction | 1 |
| II. Types of Command. | 3 |
| III. Special Command Requirements | 4 |
| IV. The Command Transmission Channel | 9 |
| 4.1 Geophysical Characterization of Convolutional Effects | 13 |
| 4.2 System Characterization of Convolutional Effects | 17 |
| 4.3 Characterization of Additive Disturbances | 20 |
| V. Error Monitor and Control Techniques | 23 |
| 5.1 One-Way Error Control Techniques | 23 |
| 5.2 Feedback Techniques | 25 |
| VI. General Signal Design Considerations | 26 |
| 6.1 Baseband and Intermediate-Band Design | 27 |
| 6.1.1 Techniques for Achieving Security | 30 |
| 6.1.2 Synthesis of Baseband Command Signals | 32 |
| 6.2 Selection of RF Modulation | 49 |
| VII. The Selection of Binary Waveforms for Digital Command Coding | 53 |
| 7.1 The Binary Detection Problem | 54 |
| 7.2 Coherent Detection of Nonfading Signals | 56 |
| 7.3 Noncoherent Detection of Nonfading Binary Signals | 60 |
| 7.4 Binary Detection in the Command-Destruct Receiver Environment | 62 |

DESIGN OF SPACE-VEHICLE COMMAND SIGNALS

I. Introduction

Command by remote control is a vital function that requires a radio communication link of a special type. In space and missile technology, it is necessary to have ground control of whether or not a mission should be completed, and to possess the capability of initiating, controlling, or blocking certain functions in the inaccessible vehicle.

Command communication, however, is but one of many communication functions that are required in a congested spectrum. Several command signals for different vehicles, or different functional blocks in the same vehicle may often necessarily overlap in time and spectrum, or both. It also appears next to impossible to secure an exclusive internationally clear radio-frequency channel either for range-safety command-destruct or for satellite and space vehicle command control. National and foreign services of various kinds occupy either the same, or contiguous, channels to those currently used for range-safety and satellite command. It is therefore necessary that the command system design take full account of the unavoidable conflicting signals and channel disturbances to ensure the degree of reliability dictated by mission requirements.

Of central importance in the planning and design of any type of communication system is the signal design problem. By signal design is meant the problem of transforming the original message into an electrical

signal that is suitable for transmission over a given medium, subject to specifications on required system performance, allowable complexity, power limitations, etc. Before one can proceed with the design of a satisfactory signal, it is of the utmost importance that the vital characteristics of the message to be transmitted be specified, that the channel through which the signal will be transmitted be well characterized, that the performance required be defined by a measurable index or set of indices, and that all of the necessary limitations on the final design (such as bandwidth, power, complexity, cost, size and weight of receivers and decoders, reaction time, etc.) be specified.

This report is intended as a survey of the types, requirements and techniques of command signal design. The material and discussions combined here is intended to provide a basis for conceiving and evaluating specific signals for special command purposes, and to provide a guide for the classification of command signals. Special examples of the various general categories are described merely for purposes of illustration and not for definitive evaluation for any special application.

II. Types of Command

It is convenient to distinguish two general types of command:

- (a) Simple, one-shot, binary-like functions, such as change-of-state function - on-off switching, closing or opening of a relay, etc. Such a function may be triggered by means of a signal occupying finite time and frequency slots, the structure of the signal being chosen to identify or address one out of several vehicles or one out of several subsystems within a vehicle.
- (b) Complex instructional-type command requiring the the flow of a sequence of simple events for the purpose of performing an elaborate maneuver or a continuous adjustment. This type of command may thus consist of a sequence of simple binary-like logical steps, starting with a one-shot change-of-state functional command.

In later sections, reference will be made to these types of command as type II(a) and type II(b).

III. Special Command Requirements

Among the special requirements that may be imposed upon a command system, three merit especial attention:

(i) Reliability, or degree of assurance with which a command signal, which is likely to be corrupted by noise and distortion, will be correctly received by a receiver whose performance is uncertain and will be properly decoded by a command decoder which is subject to the vagaries of malfunction. Thus, the reliability is determined by the probability of equipment failure and by the probability of incorrect reception or improper decoding. Reliability assurance may be secured not only by improved equipment design and coding-decoding strategies, but also by system-wide planning of appropriate verification feedback strategies.

(ii) Jam-proofness, or immunity to multi-signal environment. The desired degree of jam-proofness is limited essentially by the boundary conditions on the system design mentioned in the introduction. For example, it may be specified that the reception be maintained in the face of a CW interference that may be up to 26 db above the desired signal, and that interruptions (caused by radar signals and the like) of up to 2 milliseconds be incapable of introducing bit errors.

(iii) Security from other commands in the same time-frequency slot intended either to spoof the desired vehicle, or to command a different vehicle. The degree of necessary security depends upon the seriousness of the consequences of a false command, and the degree of assurance with

which the command function must be executed in the presence of intentional or unintentional disturbances. A quantitative measure of security may be provided in terms of the probability that any possible source of interference will succeed, either accidentally or by deliberate repeated trials, in duplicating the correct command code within the expected time duration of the mission.

In view of the requirement for continuous information flow (even though it may be to a very limited extent) in commands of type II(b), and the one-shot "trigger" nature of commands of type II(a), type II(b) commands can be expected to offer a more difficult challenge in securing specified degrees of jam-proofness and security. In particular, the considerable similarities in characteristics among transducers used for different complicated functions impose added severity to the requirement of security among subsystem commands of type II(b).

For purposes of defining the requirements for jam-proofness and security, it is convenient to distinguish between command destruct and command control. Also, space vehicles may be classified as near-earth and deep-space. In the near-earth environment, the vehicle may frequently be expected to pass over particularly hostile territory during intervals in which friendly command is not required or perhaps not advisable. More generally, the proximity of the trajectory to the earth's surface makes it possible to encounter considerable transmission path length differences to various points on earth, and therefore may place certain localities at a greater advantage than others because of shorter range to the vehicle. Under

these conditions, the desire for friendly command over certain territories may be inhibited by an overwhelming measure of transmission superiority on the side of the enemy. Only under such a condition need the requirement for security override in importance the requirement for jam-proofness. Otherwise, both of these requirements must be equally weighted in the command signal design considerations.

The many considerations that enter into the planning and design of a command communication system do not carry the same weighting in the various types of command that may be desired in a given mission starting from take-off all the way to final orbital or other operation. For example, it may not be anticipated that intelligent jamming per se will be an important consideration for sometime in the command of space laboratories and scientific and communication satellites. Security per se also may not be expected to be an important consideration. But unintentional interference and general mutual interference among various systems utilizing the 100 - 500 mc band is of vital importance both in command control as well as in range-safety command destruct. The emphasis in satellite command control may therefore be only on the ability to suppress spurious unintentional interference and the reduction of the probability of accidental command code imitation by random disturbances and by spurious cochannel and adjacent-channel signals of the kind normally encountered in a congested spectrum or arising from badly designed systems with serious RFI -- rather than signals of the type that an intelligent, knowledgeable jammer would introduce.

Emphasis in the command control situation should also be placed on low-power requirement, simplicity of receiver and decoder design and implementation, longevity of operational reliability far in excess of what is usually required in range-safety command systems, small size and weight, avoidance of the need for critical adjustment, etc.

In the range-safety command destruct application a number of considerations enter into the assessment of the requirements.

In the first place, one may expect that only a small number (perhaps 1 percent or less) of launched vehicles may have to be destroyed.

Second, the command-destruct function may only be assigned a limited life or "mission time" (say the first 5 to 8 minutes after initiation of launch) after which the critical command-destruct equipment in the vehicle may be totally disconnected or switched off from the receiver either by an automatic timer or by direct radio command.

Third, the intruder's terminal objectives that must be frustrated must be clearly identified. Two such objectives are

- (i) to block a desired command, which is essentially a jamming function,
- (ii) to trigger a false command, which essentially imposes a security requirement.

Fourth, there are factors against an intruding agency that should be taken advantage of in simplifying the required system. Among these are:

- (i) "mission time" T,

- (ii) relative locations of opponent and of range-safety officer during T,
- (iii) antenna directivity during T (save for booster exhaust plasma effects),
- (iv) knowledge of mission code.

The design of a signal to provide the required degrees of jam-proofness and security must take advantage of two flexibilities common to the design of all radio signals. The first of these is the design of the baseband signal, which amounts to an encoding of the important characteristics of the message. The second is the design of the radio-frequency, or band-pass signal, which, in turn, amounts to a transformation of the baseband signal into an r-f spectrum. It is possible to associate much of the built-in security of a signal with the baseband and intermediate-band design and the jam resistance with the r-f modulation. The r-f modulation may be selected so that it enables the use of effective interference-suppressing pre-demodulation and post-demodulation circuits. In the following sections, we review briefly the major considerations that guide the design decisions in the baseband and r-f modulation. Foremost amongst these are the characteristics of the transmission disturbances and spurious signals that may be expected at the receiver.

IV. The Command Transmission Channel

A central issue in the design of a reliable communication link is the satisfactory characterization of the transmission medium with regard to its effect upon the signals. The purpose of the characterization of the transmission disturbances is 1) to provide an aid in determining optimum signaling and receiving techniques, 2) to provide the means for determining the time and frequency resolutions achievable through a channel in terms of measurable parameters of the channel, and 3) to enable the evaluation of the degradation in information transfer due to adverse conditions between receiving and transmitting antennas.

The characterization of physical channels may be approached from a number of viewpoints. First, there is the geophysical approach which attempts to describe the propagation mechanism of the channel on the basis of applicable geometrical and physical concepts and laws. Second, there is the systems approach, which is of special interest to the communications engineer who regards the transmission medium as a "black box" that need be characterized only with regard to terminal behavior. A third approach combines a considerable portion of the terminal equipment and modems with the physical medium, and emphasizes the nature and the statistics of the errors that are encountered at the receiver.

Although currently available data is far from adequate for a quantitative characterization of the space command channel, some qualitative phenomenological descriptions can be provided that offer much needed guidance in the selection of modulation and signal-processing methods.

As with all other channels, the space channel introduces two types of disturbances that are of fundamental concern to the system designer: a signal-dependent part and a part that is independent of the signal. The signal-dependent part is termed the convolutional (or more commonly, but less accurately, multiplicative*) disturbances; the part that is independent of the signal is termed the additive disturbance. Both types are essential to a complete characterization of the channel, although the former is the part that, strictly speaking, represents the irregularities of the response of the channel at the receiving end to the signal emanated by the transmitter. The convolutional effects can be usually attributed to changing-multipath transmission and to fluctuations in the constitution and refractive properties of the intervening space. The additive perturbations result from spurious signals and disturbances that are added to the signal and may appear with and without it.

The additive noises of the space channel include all of the categories encountered in other communications:

Random-fluctuation noise mainly from the receiver front-end,

CW-type interference from spurious sources, including ground VHF and UHF links or from spurious products of front-end loading or mixing, and

Pulsed or impulsive disturbances mainly from other equipment.

* This somewhat misleading term derives from the fact that in many communication problems the non-additive effect of the channel upon the signal can usually be represented by a random multiplier.

This classification is based upon the fact that these three categories cause distinctly different interference effects, are most conveniently represented and analyzed by distinctly different analytical tools, and usually are best handled by different suppression techniques.

A much less developed subject is non-additive, or convolutional disturbances. The convolutional noise may be classified into:

- (a) Changing-multipath-interference fading. This takes the form of rapid fluctuations in the instantaneous signal strength and phase whose cause can be traced to interference among two or more slowly varying replicas of the signal arriving via different paths. This type of fading can lead to a complete loss of the message during time intervals that are long even when compared with the slowest components of the message.
- (b) Random-fluctuating diffraction, giving rise to variable path stretching, and hence variable transmission delay.
- (c) Attenuation fading caused by absorption or deflection of signal energy in the intervening medium. In addition to the usual $1/r^2$ fading caused by an increasing range, r , there may be a fluctuation in signal strength that is caused by vehicle tumbling with a non-isotropic antenna, or a much steadier severe attenuation caused by the plasma associated with the booster flame or with the heating on re-entry.

During the ascent and injection phases of a vehicle mission, booster flame attenuation plus relatively mild multipath are the chief offenders. Essentially one direct signal plus one or more weaker earth (or other) reflections reach the receiver at the launching site during the first few seconds after lift-off. Signal envelope fluctuations of the order of 10 db have been observed. Soon afterward a heavy attenuation of up to about 30 db sets in, accompanied by severe cross-polarization and noise-like fluctuations of envelope and phase. These effects are caused by the booster exhausts. At forward stations down-range, the signal strength grows steadily as the vehicle approaches the station, but mild-to-severe multipath fluctuations will be observed. There is also the possibility of fluctuations caused by yaw movements with a non-isotropic antenna. As the missile recedes from a receiving station, a change in the inclination of the missile relative to the receiving site may remove the flame plasma from the direct path to the receiver, thus suddenly lifting a very substantial amount of attenuation and disturbance and restoring the signal strength to near its free-space attenuated value. But a strong two- or multi-path mode of propagation begins to set in, caused by reflections from various surfaces (principally the earth, in the VHF and lower UHF regions) and by path-splitting in stratified and irregular atmospheric layers. The second path may start as solidly specular and turn into near-scatter. As the missile approaches the horizon and goes beyond, the received signal loses its specular character, thus becoming scatter-type.

If widely spaced ground transmitting antennas are used to illuminate the vehicle, it can be reasoned that the instantaneous fluctuations in

S/N ratio in the reception from each of the transmitting sites is almost completely independent of the instantaneous fluctuations experienced with the other signals. In other words, at times when the signal from one of the locations is observed to fade to a very low level, a similar signal from some other sufficiently distant site may very probably arrive at a much higher level compared to its own ambient noise. This means that transmissions emanating from widely separated ground control stations will in general manifest unavoidable differences in the quality of reception in the vehicle. This divergence in the quality of vehicle received signals originating at different stations can be advantageously employed to improve command-control performance in a number of situations.

4.1 Geophysical Characterization of Convolutional Effects

Multipath Effects

Transmission through multiple paths in the VHF and higher bands is largely a consequence of wave reflection from the earth's surface and from large stationary or moving objects, such as mountains, buildings, aircraft, and the like, and of path splitting and scatter by atmospheric stratifications and irregularities. Thus, in addition to the signal arriving along the direct path, reflected waves whose intensities vary with the properties of the reflecting medium and with the grazing angle are likely at all times. The magnitude and phase lag associated with the reflection coefficient vary widely with the grazing angle and with the

polarization of the incident wave with respect to the reflecting surface. For example, for reflection from a smooth sea of signals in the 200 mc band, the magnitude of the reflection coefficient remains near unity for all grazing angles with horizontal polarization, while with vertical polarization it dips below 0.6 for grazing angles in the region from 1° to about 8° .

With the likelihood of attenuation of the direct signal due to various causes, the direct and reflected signals arriving at the receiver may be arbitrarily close in amplitude. Consequently, conditions ranging from mild or no disturbance to severe multipath disturbance may be expected to arise, on a given command link.

It may be noted that much of the multipath caused by earth reflections may be reduced or eliminated by use of a highly directional antenna pattern with provisions for proper orientation. Highly directive patterns are, however, difficult to achieve at VHF and the lower UHF bands.

Atmospheric - Refractive Effects

Propagation through the earth's atmosphere subjects signals to refraction effects due to the variation of propagation phase velocity in the different layers of the earth's atmosphere. Irregular and randomly variant refraction results in two effects that are sometimes considered separately:

- (a) a bending of the ray, amounting to a lengthening of the path of propagation, and

- (b) a reduction in the group velocity resulting in an added extraneous time delay.

The total phase shifts are strongly dependent on the angle of elevation of the vehicle, because a smaller angle of elevation implies a longer distance to be traversed by the radio wave through the atmosphere. Phase shifts due to refraction can be appreciable but are quite stable in general. These phase shifts change slowly as the vehicle changes its elevation with respect to the tracking and command station, and vary with diurnal changes in the propagation velocity in the various layers of the atmosphere. They are also dependent on the frequency of the radiated waves.

These propagation effects will contribute significantly only to tracking errors. Refraction involves an apparent change in angular position which is a source of error. Refraction will also present a slightly changed vehicle aspect angle which will cause an erroneous component of vehicle velocity to be observed. Range measurement errors deserve particular attention. A ranging system actually measures the propagation time delay over a particular path length. Thus the reduction in group velocity will cause a range error if the propagation velocity must be assumed constant (a velocity error will also arise from this same effect) and range errors will arise from path length variations.

Plasma Effects

Radio waves may be absorbed, diffracted and reflected by the plasmas that form in the booster flame region during the powered phase of the flight, and around a hypersonic vehicle during the re-entry phase. In response to an applied electromagnetic field, the vibrating electrons in the plasma act as parasitic radiators alternately absorbing and re-radiating energy. Collisions introduce a randomizing effect that keeps the excited plasma from re-radiating coherently with the applied field. The net effect of this randomization is to cause a signal attenuation, signal frequency spread, spurious modulations, loss of phase coherence between signals, and a change of polarization. Plasmas also strongly affect antenna matching characteristics, radiation patterns, and power handling capability, and reduce the field strength necessary for breakdown in the aperture region. Breakdown is a condition in which the air surrounding the antenna aperture becomes a strongly conducting medium that will not support wave propagation. Antenna breakdown is a very common cause of telemetry, command and communication dropout.

This disruption may only be an inconvenience if real-time communication is not required. For example, telemetry signals can be recorded on-board for retransmission after the plasma effects diminish, or they may be recorded in a recoverable instrument package. Unfortunately these techniques are not feasible if a "real-time" continuous

flow of command or other information is required. Reliable communication links must be designed for manned re-entry vehicle missions in which continuous ground-to-vehicle communication is necessary during all phases of the flight.

4.2 System Characterization of Convolutional Effects

From the viewpoint of the communication system designer, satisfactory characterization of the channel to be employed requires:

1. The definition of parameters that characterize the effects of the channel upon broad classes of signals, and thus specify a channel model that is complete for purposes of system design.
2. The determination of interrelationships among the channel parameters.
3. The determination of techniques for measuring and/or estimating the parameters of particular channels.

The system designer is mainly concerned with excitations and responses and the fidelity with which a given excitation is reproduced in the response, where fidelity may be interpreted in a number of ways depending upon the situation at hand. In the expression of the responses in terms of the excitations, especially when dealing with transmission media that can be considered to be linear, the system designer relies heavily upon mathematical representations of the signals in terms of time as the independent variable, or in terms of frequency. In either case, the signal

is conceived to be made up of a sum of appropriately weighted building blocks and the assumed linearity of the medium allows one to add up the responses to the individual building blocks to obtain an expression for the response. The transmission medium is considered to modify each of the building blocks in a manner that is characteristic of the medium. The function of time or frequency (or both) that describes the manner in which the system modifies the building blocks employed in the representation of the input signal in order to yield the response is characteristic of the system and it may be called the characteristic system function of the medium.

The fundamental philosophy of the system approach to the characterization of a transmission medium is to represent the medium by a characteristic system function. By definition, such a function relates the response and the excitation of the channel in a unique manner, and is therefore completely representative of the channel as observed at specified input and output pairs of terminals. The time, frequency and/or statistical behavior of a characteristic system function offers the basis for the definition of gross parameters for describing average aspects of the channel response.

Transmission channels can (with few exceptions) be represented generally as randomly time-variant linear filters. The characteristics of such filters should clarify the limitations the channels impose on signaling through them.

The most general randomly time-variant linear channel is one that may be defined as a linear channel whose response to a deterministic signal consists of the sum of a random component and a deterministic component. Virtually all communication channels may be regarded as special cases of this general channel. From a conceptual, or even physical viewpoint, it is convenient to regard this channel as the parallel combination of two time-variant linear filters, one random and one deterministic, followed by the addition of a noise, as illustrated in Fig. 1. The deterministic filter provides the deterministic component of the received waveform and the random filter provides the convolutional disturbance or that portion of the random component of the received waveform which is signal-dependent, (i. e., which vanishes when the transmitted signal vanishes). Finally, the operation of adding noise is used to represent the additive component of the disturbances experienced at the receiver (i. e., the disturbances that persist in the absence of transmitted signal).

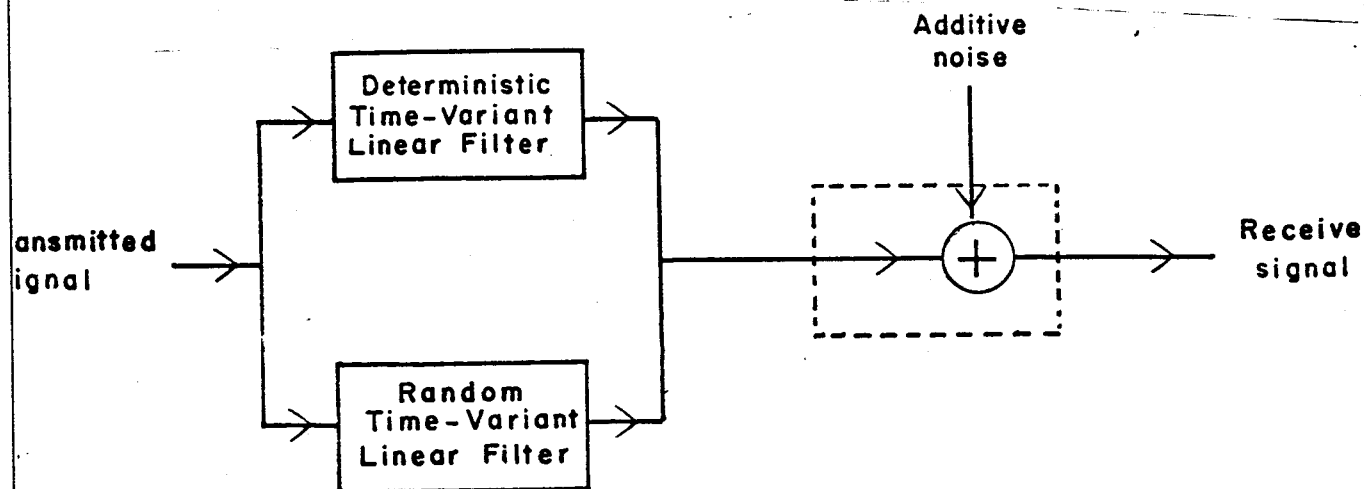


Fig. 1 Representation of a Random Time-Variant Linear Channel

Thus, for a space communication or command link, a knowledge of tracking data and attenuation characteristics of the propagation medium provides a continuous estimate of the magnitude and phase of the deterministic component as it changes slowly with the trajectory of the vehicle. The multipath and fluctuating attenuation uncertainties and errors of this estimate relative to the actual received signal, plus the additive random noise, represent the random signal component.

4.3 Characterization of Additive Disturbances

One of the first additive disturbances that is always a factor in performance evaluation is noise generated in the receiver front end. Providing receivers with adequate sensitivity may be achieved with "state of the art" techniques and components. As a result, most receivers can be designed to meet the necessary sensitivity requirements demanded by a given missile and satellite trajectory. The problem of guaranteeing reliability of signal reception often does not lie so much with sensitivity. The strain on maintaining reliability is often caused by the presence of other signals that corrupt the spectral region of the desired signal.

The command r-f signal will probably lie in the upper VHF or lower UHF range (i. e., in the 100 to 500 mc range). A channel assignment may perhaps consist of allocating the center frequency plus a 1 mc guard band about this center frequency. The specifications may also

require that the receiver selectivity characteristic have 60 db rejection at ± 0.5 mc from the center frequency assignment. It may be expected that this frequency channel assignment will not conflict with other non-command assignments, but the same r-f channel may have to be used simultaneously for commanding several vehicles distinguished only by their different address codes. However, in the 406-450 mc range alone there are at present about 500 active 600 watt and 10 kw FRW-2 transmitters, plus other pulsed CW radars in addition to numerous navigational radars assigned to this frequency range. Thus a specific assignment of frequency does not guarantee exclusive rights to that frequency and associated channels. When a missile is launched, it is illuminated by a large number of radar transmitters that were not or could not be accounted for prior to the launch. An orbiting satellite is often subjected to the same treatment. These interfering signals may originate from a foreign-based transmitter, a nearby ship or from a radar network such as the DEW line or other early warning systems.

Enemy jamming or deliberate attempts to spoof the command communication link do not recognize any boundaries in the frequency spectrum. But this type of disturbance may or may not be of significance, depending upon the mission. These deliberate jamming signals may take any form. The simplest would be a continuous transmission

of a simple sinusoid. This signal alone is enough to jam an FM system should it (or a related intermodulation term generated by some non-linear front-end process) appear in band with the desired signal and have sufficient magnitude. The jamming signal may also be a frequency-modulated signal that sweeps across the spectral region containing the desired signal. The modulated sweep signal may take the form of some pseudo-random FSK that causes a high-frequency carrier to hop around the band. The jamming signal may also consist of high level random noise flooding the band. Such a situation again raises sensitivity problems for the receiver and decoder even though the signal may be well above the front end noise level.

Along with CW interference or random noise interference, there are other interference effects generated by radar signals. These arise from discontinuities in the interfering signal. For example, pulsed CW radars have envelope modulations that undergo very abrupt changes in signal level. These are difficult forms of interference to suppress in the receiver because of the transient nature of the excitation, but their effects can be abated by appropriate coding-decoding strategies.

V. Error Monitor and Control Techniques

Automatic corrective measures form an integral part of a communication system whose performance must meet the requirements of a specified criterion with a high degree of reliability. There are two classes of error control techniques.

5.1 One-Way Error Control Techniques

(i) Interference suppression techniques

These are receiver signal processing operations that attempt to capitalize on distinguishing characteristics of various co-channel and adjacent channel signals in order to suppress the undesired signals and improve the reception of the desired signal. These techniques are particularly applicable to the command reception problem.

(ii) Diversity

Diversity combining is an effective error control technique to counteract fading. The application of diversity techniques to the space command problem is limited by the fact that (a) vehicle dimensions usually do not allow much separation between antennas intended for receiving suitable diversity signals; (b) signals transmitted within the allowed frequency space do not provide noticeable frequency diversity possibilities; (c) the requirements on reaction time do not usually provide sufficient room for time diversity to be applicable; and (d) the available data is insufficient for an assessment of the applicability of polarization diversity.

(iii) Channel-estimating Adaptive Receivers

Diversity techniques are usually planned on the basis of known or measured statistically regular parameters of the channel. Recent developments have opened the way for a class of techniques that perform channel estimations and predictions of impending behavior on the basis of behavior just experienced, and automatically adapt the receiver to the requirements of the impending behavior. At the present time, these techniques remain much too complex to offer significant promise for the space command application.

(iv) Coding and Decoding

Known one-way coding-decoding techniques can be used effectively to reduce error rate when bit errors occur independently and with low frequency. However, when the bit errors are not low, these techniques become impractical. The limitation here is not in the encoding, but rather in the considerable complexity of error correctors required to correct large numbers of errors.

Burst error-correcting schemes have been conceived that can correct any pattern of errors that lie within a fixed burst length. But, the design of equipment to correct long bursts is impractical. Again, while the encoder may be relatively simple and straightforward for large classes of codes, the complexity of the required decoding equipment is usually prohibitive. However, the characteristics of the

errors anticipated in the command environment offer bright prospects for effective and simple coding and decoding schemes.

5.2 Feedback Techniques

When it is desired to communicate essentially non-redundant data with negligible error over channels subject to fading and burst errors, the only practical solution sometimes appears to be two-way rather than one-way error correction. Considerable work has been done to determine the performance of feedback error-correction schemes under the assumption of independent bit errors. Feedback error correction over links with fading and burst noise has not been studied yet. Moreover, present feedback techniques incur a loss of transmission time that is of the order of the two-way transmission time whenever an error is detected in a block. Ways for eliminating this waste of transmission time have not been fully explored. Feedback, however, does provide important prospects for verification of proper reception and processing of a command signal when the requirements on reaction time are not prohibitive. Verification is obviously not needed in range safety applications, and is often provided by the cessation of some observable activity in other situations. But there is a clear need for some suitable means of verification in many command control situations.

VI. General Signal Design Considerations

Signal design is intended here to embrace the problem of expressing the desired message in the form that is most suitable for transmission and ultimate processing at the receiving end, within all of the applicable boundary conditions and limitations. Thus, the design of the signal that will carry the message is influenced by almost every important limitation that may be imposed upon the contemplated system; e.g., limitations imposed by the behavior of the channel, peak or average power limitations, bandwidth limitations, limitations on reaction time, economic limitations, and so on. It would be improper to consider complexity and economics as issues external to the communication problem. "State-of-the-art" of realizing various functions, functional modules and subsystems more often than not place vital limitations on the feasibility of the system.

In general, the aim of signal design for channels of specified parameters and additive disturbances is to enable the receiver to extract the intended message to within the specified performance tolerances. The first aim of signal design for unknown channels is to establish some measure of successful contact using a receiver that is adaptable to a wide class of electromagnetic environmental conditions.

Signal design involves two major steps. The first is the design of the baseband and/or intermediate-band signal, which amounts to an encoding of the important characteristics of the message. The second

is the design of the radio-frequency, or bandpass signal, which amounts to a transformation of the baseband signal into an r-f spectrum. The design of a command signal to provide the required degrees of jam-proofness and security must take advantage of the flexibilities available in these steps. Let us review briefly the major considerations that guide the design decisions in each step.

6.1 Baseband and Intermediate-Band Design

In general, it is possible to consider the problem of security as being fundamentally a "baseband" signal design problem. In the final analysis, the radiated signal is always a "narrow-band" signal in the sense that its total band occupancy is a small fraction of any frequency contained within the band.

The structure of the baseband signal is determined mainly by considerations of:

- (a) Message representation, including the definition of the significant characteristics of the message that must be preserved, and the transformation of these characteristics by some encoding process into a form that possesses one or more of the following properties:
 - (i) error detecting and/or correcting features;
 - (ii) a desired degree of security through appropriate encryption;

- (iii) address information for identifying a specific message destination.
- (b) Added security and jam-resistance through the use of noise-like, or other, subcarriers to derive an intermediate band in which the message is deliberately masked by the 'noise-like, or pseudo-random spread-spectrum subcarrier.
- (c) Multiplexing. The necessity for transmitting in real time a number of messages having overlapping spectra and times of occurrence makes multiplexing inevitable. The standard multiplexing methods are those commonly classified as frequency-division and time-division. In general, however, multiplexing of similar messages can be achieved:
 - (i) by shifting the original spectra (or spectra derived in a unique manner from them, as in FM) of the messages into different non-overlapping frequency locations (i.e., FDM),
 - (ii) by sampling the different messages during different non-overlapping time apertures (i.e., TDM),
 - (iii) by modulating orthogonal carriers (or subcarriers) other than sine waves of different frequencies and non-overlapping pulse trains, or

- (iv) by some special technique, such as "power-division" multiplexing in FM.

In frequency-division multiplexing (FDM), a composite baseband is usually synthesized by linear-modulation (AM, SSB, DSB, etc.) techniques, or by exponential-modulation (FM, PM). The composite baseband may then be used to generate a radio-frequency signal, again by linear modulation or by exponential modulation.

The design of the baseband spectrum and the choice of "subcarrier" modulation in synthesizing the composite baseband is primarily based upon the following considerations:

- (a) simplicity of generation and, to a much greater extent, of demodulation,
- (b) spectral properties of the time functions used in the command encoding,
- (c) bandwidth requirements of the "subcarrier" modulation method,
- (d) sensitivity of resulting multiplexed spectra to noise and distortion, and
- (e) availability of appropriate well-developed equipment, either in a current investment or on the market.

In the design of the baseband waveforms, it is found that techniques that alleviate the undesirable effects of one kind of noise often

increase the undesirable effects of another. Thus channel equalization may reduce intersymbol interference but it may at the same time increase the effect of additive noise at the output. As another example, the narrowing of the pulses of a particular channel reduces intersymbol interference but increases adjacent channel interference. It is thus clearly of interest to know what combinations of pulse shapes and receiver filtering are optimal from the viewpoint of reducing the combined disturbance due to additive noise, interchannel, and intersymbol interference.

6.1.1 Techniques for Achieving Security

The design of the baseband signal may employ two general approaches:

(a) Special coding in terms of:

- (i) Tone Sequences: For example, a certain sequence of tones all emanated at the same time, or at specially precoded times, whose frequencies bear a certain pre-arranged relationship (and all may jump in a certain pre-arranged manner at exactly the same instant of time through especially coded frequency jumps).
- (ii) Pulse sequences
- (iii) Pulsed tones

(b) Spread spectrum

(i) Time/Frequency hopping

(ii) Pseudo-random sequences

In spread-spectrum techniques, every information symbol is expressed in a waveform code that carries the address of the intended system or subsystem; i.e., identifies the desired receiver. The receiver is provided with a decoder that responds with high probability only to the signals addressed to it.

Pseudo-random sequences and time functions are so named because, although they are generated deterministically by means of a code of long period, they appear to be random to an observer who is ignorant of the code and who has available only a segment of the sequence.

Large duration-bandwidth product signals are important in the design of secure and/or jam-resistant systems because:

(a) A large symbol duration

- localizes impulsive and short-pulsed disturbances within a small fraction of the symbol duration, and hence reduces the effectiveness of such a disturbance,
- samples random noise over such a long duration as to enable an effective smoothing of it,
- provides a high degree of flexibility for elaborate coding of the desired signal.

(b) A large bandwidth compared with the width of the desired baseband

--localizes narrow-band disturbances severely in frequency, thus enabling the receiver operations to spread them out over the much wider transmission bandwidth and to filter out the substantial fraction of the spread out spectrum that falls outside of the desired information bandwidth,
--forces the enemy to increase his transmitter power and to spread his power effectively over a wide range.

6.1.2 Synthesis of Baseband Command Signals

Word format typically includes: vehicle address, address for subsystem in a vehicle, desired command message, block. The vehicle command address should uniquely identify the desired vehicle, leaving all other unintended vehicles unresponsive. Similarly for the subsystem address. Provisions may also be desired for triggering a signal confirming proper reception, identification, delivery and execution of a command. The various parts of a complete command word may be synthesized in a number of ways. For example:

(i) Simple tone techniques.

1. One tone carefully selected not to correspond to a harmonic or a beat between anticipated components. Vanguard used one tone.

2. A number of such tones, turned on simultaneously.
Up to seven tones have been used.
3. A number of such tones turned on for different
partially overlapping time intervals.

(ii) Modulated subcarrier techniques.

1. In one technique:

One 50 kc subcarrier is frequency-shift modulated ± 15 kc. The modulated subcarrier in turn frequency modulates an r-f carrier centered within one of 44 one-mc channels in the range $406 \text{ mc} \leq f \leq 450 \text{ mc}$. The frequency deviation of the r-f carrier is ± 40 kc.

2. The Gemini digital command system, uses:

--2 kc subcarrier phase-reversal keyed by command data at 1000 sub-bits/sec (2 cycles of subcarrier per sub-bit);

--1 kc subcarrier conveys bit synchronization and phase reference for optimal phase comparison in the detection process.

3. The Goddard Tone-Digital System

In this system, PCM coding amplitude modulates audio subcarrier (or tone). Tones are selected

in the band from 7000 cps to 11,024 cps. The signaling strategy is as follows:

- Unique address word of eight bits is repeated twice
- Function command word of 8 bits is repeated 3 times
- Correct reception of one correct address command word plus one correct function command word in the same sequence of 5 words is sufficient to effect a command.
- In a valid detection sequence:
Synch pulse is detected first
then once the address has been detected, a valid function command word must be read within the time interval for the 5 word series of 2 address command and 3 function command words.

The code words are synthesized from a prespecified number of zeros and ones.

Address command word: 3 ones and 5 zeros, or,
5 ones and 3 zeros.

Function command word: 4 ones and 4 zeros.

This coding provides 112 unique address command words, so that the same tone can be employed for

112 satellites. Total number of satellites that can be handled is 112 times the number of available tones.

The 4×4 function command word format provides 70 different commands. Because this many commands is not usually needed, provisions for 6-bit words (with 2 ones and 4 zeros or 4 ones and 2 zeros for address, and 3 ones and 3 zeros for function) are made in ground station design. This provides 20 commands.

4. The Goddard Address-Execute Tone System

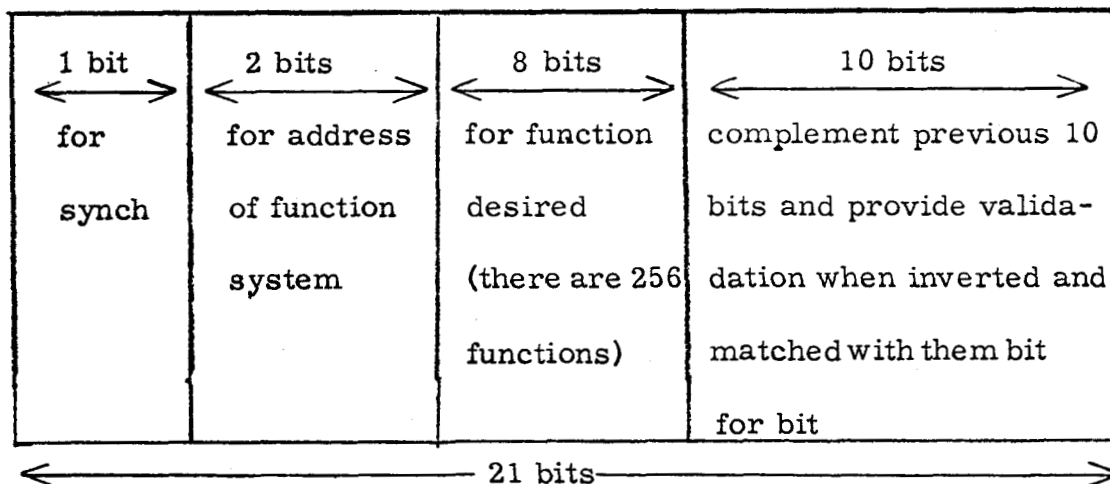
This system is intended for handling up to 15 satellites each requiring less than seven commands, with some immunity against extraneous false commands and protection against accidental command by a NASA station.

The signaling strategy is to send a serial sequence of an address tone (unique to a satellite) followed by one or more execute tones. Address tone turns on satellite for a preset interval long enough to receive the execute tones. Command decoder requires the tones to be present for a minimum length of time (0.5, 1.0, 1.5, 2.0, 2.5, or 3.0 seconds) before it registers the command, rejecting tones

of shorter duration. Twenty-two tones are selected in the range of 1025-6177 cps. The 9 tones in the first octave and the 6 tones in the 3rd octave are used as unique address tones. The 7 tones in the 2nd octave are for execute tones.

5. Instruction Command Systems

--OGO satellite uses a 21-bit command word sent at a rate of 128 bits/sec. Data-word carries both word synch and bit synch. Binary FSK between 2 tones is used. The composition of the complete command signal is as follows:



--The OAO commands are composed of two 32-bit words. Each word is accompanied with a transmission of its complement. The bit rate is 1000 bits/sec. Frequency shift keying between two

tones is employed for the data word. Validation occurs by on-board checking of a word against its complement and this is verified on the ground by echoing the validated word to the ground. In all but the synch commands, the first word gives information to the satellite about what to do with the second word. The first two bits are used for word synch. The third bit determines whether this is a real-time or a storable command. In stored commands, bits 4-13 give the execution time and bits 14-20 give the memory location for storing both words. For both real-time and stored commands, bits 21-24 indicate the type of command.

(iii) Time/Frequency tone hopping techniques

In time/frequency hopping the allotted time for a one-bit command (say 50 msec) is subdivided into N subintervals of equal duration (say 5 msec) and each subinterval is used for transmitting a "sub-bit" in the form of one or more simultaneously pulsed tones whose frequencies are chosen from a discrete set in some specified range (say 7 to 20 kc).

Special rules must usually be observed in the selection of the set of acceptable tone frequencies, especially if the

pulsed tones are conveyed by modulating the phase or frequency of an r-f carrier, or if nonlinearities may be expected in the transmission and processing of the pulsed tones. For example, if confusion and spurious false sub-bits are to be minimized, it is desirable to choose the acceptable frequencies to satisfy the following requirements:

- (a) No frequency should be a harmonic of another.
- (b) No beat between two or more of the frequencies should, either itself or a harmonic of it, be equal to one of the acceptable tones.
- (c) For decoder simplicity without considerable sacrifice of security, choice of a small number of frequencies (say seven) would be preferable, with allowance for up to two frequencies to be transmitted in each time slot.
- (d) Lower frequencies (7 - 20 kc) are preferable, perhaps falling between IRIG command channel assignments (see Table I).
- (e) Each frequency f_0 should ideally be chosen so that none of the others or their harmonics would fall within $f_0 \pm \Delta F$ cps, where ΔF denotes one-half of a suitably defined frequency resolution bandwidth.

TABLE I

Tone Frequencies of 20 IRIG Command Channels (1 through 20)

| <u>Channel</u> | <u>Frequency, kc</u> |
|----------------|----------------------|
| 1 | 7.5 |
| 2 | 8.46 |
| 3 | 9.54 |
| 4 | 10.76 |
| 5 | 12.14 |
| 6 | 13.70 |
| 7 | 15.45 |
| 8 | 17.43 |
| 9 | 19.66 |
| 10 | 22.17 |
| 11 | 25.01 |
| 12 | 28.21 |
| 13 | 31.83 |
| 14 | 35.90 |
| 15 | 40.49 |
| 16 | 45.68 |
| 17 | 51.52 |
| 18 | 58.12 |
| 19 | 65.56 |
| 20 | 73.95 |

In one time/frequency coding scheme, the baseband format consists of a total of 10 bits that constitute an 8-bit address followed by either an ON or OFF bit and a Function bit. Each bit is conveyed by a tone of 5 msec duration, selected in any desired order from a set of 19 discrete frequencies in the range of 7 to 13 kc. The frequencies and pulse waveforms of the tones ideally are chosen so as to minimize the probability of false bits resulting from nonlinear treatment of the tones (in modulation, demodulation and signal processing) and to minimize intersymbol interference resulting from time overlaps. The pulsed tones are detected by matched filters followed by decision circuits with decision thresholds set to minimize the probability of decision error. The bit duration (5 msec) is chosen to be considerably longer than the duration (2 msec) of the longest expected disruptive radar pulse in order to prevent such pulses from masking the command bits.

The pulse shape of subcarrier tones must also be chosen with care. Two conflicting requirements influence the selection of pulse shape for the subcarrier tones.

First, irrespective of the r-f modulation, the pulses in adjacent time slots should ideally not overlap at all. From the viewpoint of anti-jamming capability, the pulse amplitude must remain at the maximum possible level during its entire nominal duration. Thus, if the pulses are transmitted on a wideband FM carrier, it is then desirable that the peak factor of the instantaneous frequency deviation of the r-f signal be as low as possible. In this way, the spectrum of the r-f signal would be most effectively spread out, and the wideband advantages of the FM signal would be fully realized. At baseband, the energy packed in the pulse must ideally be at the maximum level allowed by the available power. The pulse shape that meets these requirements is ideally rectangular.

But from the viewpoint of spectral occupancy, the spectrum of the pulse should ideally be rectangular. This avoids spectral overlap. Spectrum of the pulse is of critical importance in the selection of the frequency spacing between the baseband tones. The spectrum of the pulse itself (centered about the frequency of the tone) is not an important factor with respect to the r-f spectral occupancy because in order to survive long (say 2 msec) blanking

pulses or disturbances, the duration ($T = 5$ msec) of the desired pulse must be such that $1/T \ll$ frequency of tones selected from the range of 7 kc - 20 kc. Therefore, if two tones f_1 and f_2 are used simultaneously, spectral lines will be present at $nf_1 \pm mf_2$ away from the carrier frequency in the FM spectrum and these spectral components, rather than the individual pulse spectrum, will have the major effect upon the r-f spectral occupancy.

The ideal (rectangular) pulse shape that meets the requirement for maximum A/J capability incurs a $|\sin x/x|$ spectrum that limits the frequency spacing of the tones severely. The other extreme pulse shape that ideally meets the requirement for restricting the spectrum of the pulse has a $|\sin x/x|$ shape that limits the time spacing of the tones severely. A compromise must therefore be made in terms of a waveform that is compact in both time duration and spectral occupancy.

There are two pulse shapes that are of special interest in the search for a high degree of time and frequency "compactness." These are the gaussian-shaped pulse, and the "cosine-squared" or "raised cosine" pulse. Each of these pulse shapes is well localized in time to within a reasonably well-defined duration around $t = \tau_0$, and their spectra are

well-localized in frequency to within a well-defined frequency band around $\omega = \omega_0$. An important consequence of this property is that inter-symbol interference arising from "spilling over" adjacent pulses, and spurious responses in the wrong filters arising from overlapping spectra are all minimized.

Gaussian-shaped pulses are generated by exciting a "gaussian" filter with a very short pulse (of almost arbitrary shape). The term "gaussian filter" is used in reference to a filter whose system function is given by

$$H(j\omega) = e^{-\left(\frac{\omega - \omega_0}{\alpha}\right)^2} \cdot e^{-j(\omega - \omega_0)\tau_0} \quad (1)$$

which is seen to have a gaussian-shaped amplitude characteristic and a linear phase characteristic. Such a filter has a gaussian-shaped impulse response also, described by

$$h(t) = e^{-(\alpha/2)^2(t - \tau_0)^2} \cdot e^{j\omega_0 t} \quad (2)$$

The half-power bandwidth can be computed by setting

$$e^{-\left(\frac{\omega - \omega_0}{\alpha}\right)^2} = 1/\sqrt{2} \quad (3)$$

whence

$$\begin{aligned} (BW) &= (\sqrt{2\ln 2}) \alpha \\ &\approx 1.178 \alpha \text{ rad/sec} \end{aligned} \quad (4)$$

Note also that $|h(t)|$ drops to $1/e$ of its peak value at

$$|t - \tau_0| = 2/\alpha \text{ sec} \quad (5)$$

and to $1/100$ of the peak value at

$$|t - \tau_0| \approx 4.30/\alpha \text{ sec} \quad (6)$$

The gaussian function (which describes the shape of both the envelope of the impulse response as well as the amplitude vs. frequency selectivity of the gaussian filter) has some very interesting properties. In order that the discussion apply both to $|H(j\omega)|$ and to $|h(t)|$ of Eq. (1) and (2), let us discuss

$$f(x) = e^{-x^2} \quad (7)$$

where

$$x = \frac{\omega - \omega_0}{\alpha}, \text{ if } f \equiv H \quad (8)$$

$$= \frac{\alpha}{2} (t - \tau_0), \text{ if } f \equiv h \quad (9)$$

Since x occurs only in an even power in $f(x)$, the first and higher-order odd derivatives of $f(x)$ are all zero at $x=0$. Moreover, in the vicinity of $x=0$,

$$f(x) \approx 1 - x^2 + \frac{1}{2!} x^4$$

$$\approx 1 \text{ for } x \leq 0.31 \quad (10)$$

These observations indicate a somewhat flat top for the variation with x . This flat top is associated with a rather sharp cut-off characteristic, as may be appreciated by noting that $f(x)$ drops from -3 db at $x = \pm 0.588$ to -20 db at $x = \pm 1.15$, to -40 db at ± 2.14 .

A db plot of $f(x)$ yields a parabola centered about $x = 0$.

Note also that the product of a number of $f(x)$ functions of the form (7) is also of the same form. This means that the overall selectivity characteristic of a number of gaussian filters in cascade is also gaussian-shaped. Moreover, the square of the magnitude of $H(j\omega)$ is also gaussian shaped.

$$|H(j\omega)|^2 = e^{-\left(\frac{\omega - \omega_0}{\alpha/\sqrt{2}}\right)^2} \quad (11)$$

The gaussian function is a useful computational tool in noise analysis (especially in computations involving mean-square spectral densities, etc.) because integrals containing functions of the form of e^{-x^2} in the integrand are either very manageable or well tabulated, or both.

Of immediate interest to the practical realization problem is the fact that if we recall the basic definition of e^a as a limit, we can write

$$|H(j\omega)|^2 = \lim_{n \rightarrow \infty} \left[1 + \left(\frac{\omega - \omega_0}{\sqrt{n/2} \alpha} \right)^2 \right]^{-n} \quad (12)$$

This shows that a practical approximation to a gaussian filter may be realized by cascading n synchronously tuned single-pole, high-Q circuits, each with a half-power bandwidth of $\sqrt{2n} \alpha$ rad/sec.

The "cosine squared," or "raised cosine" function may be expressed as

$$\begin{aligned} p(t) &= 1 + \cos \omega_p t, & -\pi/\omega_p \leq t \leq \pi/\omega_p \\ &= 0, & |t| \geq \pi/\omega_p \end{aligned} \quad (13)$$

Its Fourier transform is

$$\begin{aligned} P(j\omega) &= \int_{-\pi/\omega_p}^{\pi/\omega_p} p(t) e^{-j\omega t} dt \\ &= \frac{2\pi/\omega_p}{1 - \frac{\omega^2}{\omega_p^2}} \cdot \frac{\sin(\pi\omega/\omega_p)}{(\pi\omega/\omega_p)} \end{aligned} \quad (14)$$

The transform of $f(t)$ is a $(\sin x/x)$ -shaped function weighted by $1/(1 - \omega^2/\omega_p^2)$. As $|\omega|$ is increased from

zero, $P(j\omega)$ drops first slowly, then steeply toward its first zeros at

$$\omega = \pm \omega_p \quad (14)$$

and it remains small for $|\omega| > \omega_p$. For practical purposes, the spectrum of the pulse may be considered to be contained within $-\omega_p < \omega < \omega_p$.

Figure 2 shows a plot of the spectra of a gaussian-shaped pulse and of a raise-cosined pulse for comparison.

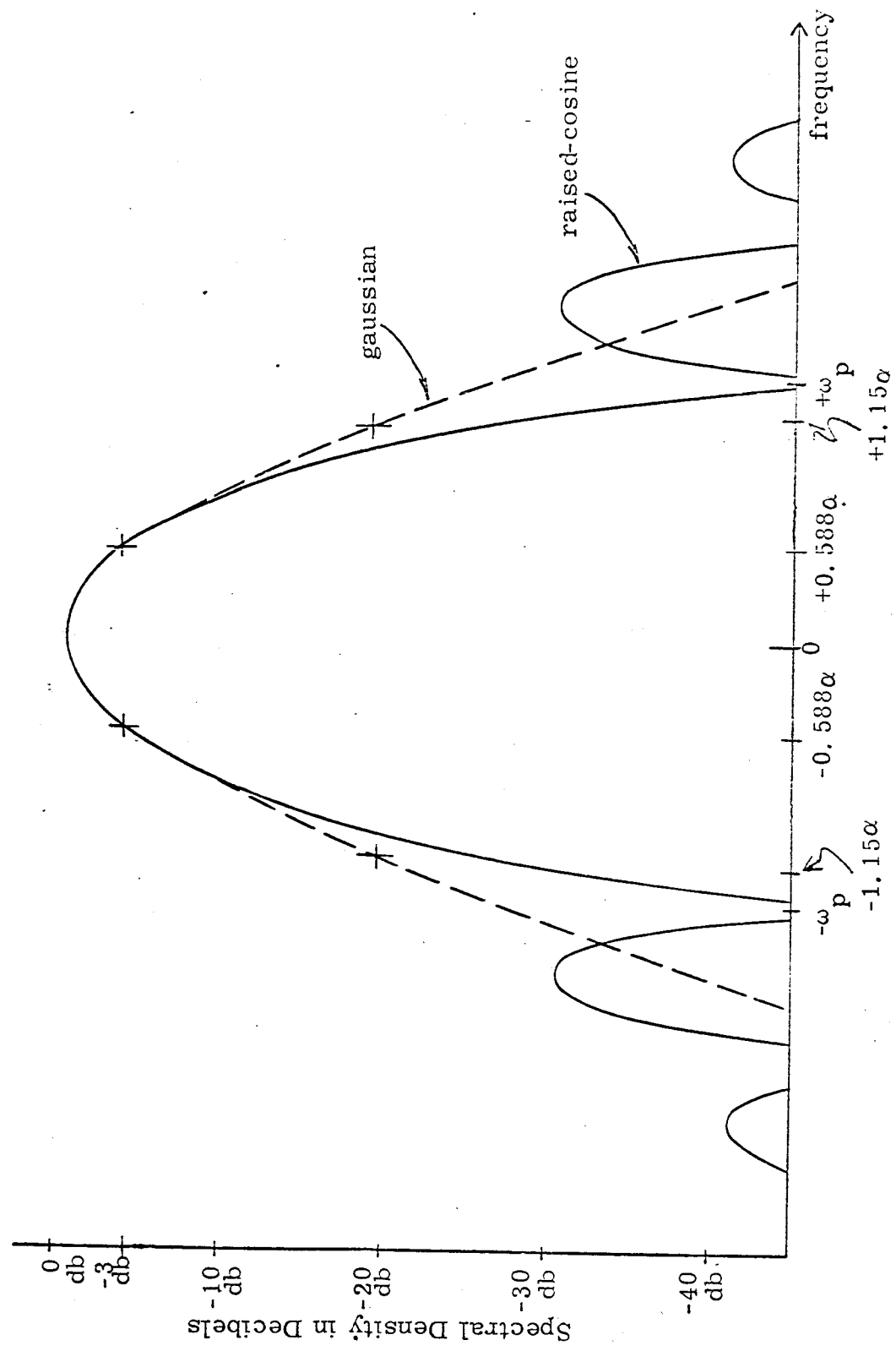


Fig. 2 Spectra of gaussian and raised-cosine pulses

6.2 Selection of RF Modulation

The important considerations in the selection of r-f modulation method are:

- (a) simplicity of system implementation, especially in the vehicle-borne receiver,
- (b) efficiency of available power utilization - specifically:
 - (i) maximization of power "packing" into the radiated signal; and
 - (ii) maximization of power-amplifier efficiency;
- (c) minimization of bandwidth occupancy;
- (d) compatibility with other types of signals employed for space operations and communications signals; and
- (e) sensitivity to expected multiplicative and additive disturbances, and amenability to simple and effective corrective measures against intentional and unintentional interference (or jamming).

These considerations may be justified briefly as follows:

Overall system simplicity is desirable not only because of cost considerations, but also because complexity usually means more functional operations to be performed requiring more circuits and more components and hence greater chances of malfunction, or lower operational reliability. It should be possible to extract the desired information with

a minimum of signal processing, thus minimizing potential sources of error and distortion in the circuitry, and maximizing performance reliability.

The question of efficiency of available power utilization arises from the fact that a limited amount of power is usually available in the command transmitter for investment in the structure of the signal. The efficiency of available power "packing" into the signal structure is a function of the peak factor (i.e., the ratio of peak value to rms value) of the envelope of the composite signal. The importance of the peak factor of signal envelope rests upon the fact that the ultimate performance of a communication system is almost always determined by average signal power, taken over a typical message interval, while peak of average power taken over many carrier cycles, is a limiting transmitter design factor in almost all practical cases of interest. In addition, low peak factor implies less stringent requirements on linearity and, therefore, improved chances for utilizing efficient nonlinear power amplifiers (i.e., Class C amplifiers).

The interest in the signal bandwidth requirements is based upon the desirability of minimizing noise bandwidth, and improving the efficiency of spectrum utilization. The former bears strongly upon range of reliable contact; the latter is motivated by the rapidly increasing number of desired communications and the rapidly dwindling

available spectrum space. Arguments for and against reduction of bandwidth occupancy can be presented. For example, lower signal bandwidths obviate the need for wideband transmission filters that usually present a greater possibility of drifting characteristics and deviations from uniformity and hence increased chances of attendant signal distortion. On the other hand, exponent modulation (e.g., FM and PM) trades increased jam-resistance and better overall system performance for wider bandwidth occupancy.

As for sensitivity to disturbances, recall from Section IV that the r-f signal will be subjected to a number of severe effects, such as strong attenuation, multipath, spurious interferences and deliberate jamming. It is, therefore, important that the type of r-f modulation used should yield a signal that is as insensitive to these disturbances as possible, especially by virtue of its amenability to corrective processing, interference suppression techniques, etc. Moreover, in the design of the baseband signals, one should take into account the effect of the r-f demodulation process upon the spectral distribution of the attendant noise and distortion over the composite baseband. For example, it is important to note that the demodulation process in some systems (e.g., FM) will alter the noise power spectral density in such a way as to favor certain frequencies over others.

It is of interest to observe that if time/frequency hopped tones are transmitted on an FM carrier, then the fact that the output FM spectrum is parabolic for $(S/N)_{in} > 12$ db, need not require that the desired tone energies be scaled up with frequency in order to maintain the same S/N, for all tones. Note that $(S/N)_{in} > 12$ db assures good tone detection. On the other hand, at and below $(S/N) \approx 9$ db the spectrum of the noise spreads out with frequency and looks almost uniform.

Generation of the r-f signal by linear-modulation techniques may result in AM, DSB or SSB. Generation by exponential-modulation methods results in FM, or PM. In instances like discrete two-phase keying, the phase-modulated signal can also be viewed as a linear-modulation signal, which simplifies the solution of several associated problems.

In the light of the above categorization of the signal design problem, the command receiver may be subdivided into two major functional blocks. The first is what may be called the "command tuner." It consists of that part of the receiver that recovers the baseband signal from the incoming r-f signal. The second part is what may be called the "command decoder." It consists of that part of the receiver that extracts the command instructions from the demodulated baseband output of the command tuner. Above all, the command decoder must embody the highest possible degree of simplicity that is consistent with the required reliability with which the decoding function must be executed.

VII. The Selection of Binary Waveforms for Digital Command Coding

In any situation, the variety of different command messages to be transmitted may be arbitrarily designated by numbers. For digital transmission purposes, these message designations may be expressed in either the binary or a p'ary number system. It is important to note that writing the message number in p'ary form is but one way of assigning a p'ary number (i.e., a sequence of p'ary digits) to each code message. Other binary or higher-radix coding methods exist which require "code lengths" or numbers of digits that may be different for different messages.

Alternately, a more elaborate designation scheme may be used in order to secure some advantages such as error correction. For this purpose a coding scheme might be desired which anticipates certain classes of errors and utilizes a designation scheme that allows certain numbers of these errors to occur in various parts of the code word without causing false recognition of the desired command. One may conceive of this operation as one in which more than one representation is given to each command and these alternate representations considered as a group of "equivalent" code words. A complete description of the coded command message set in any time interval is provided by the number of identifiable groups, or representative prototypes, plus the probability with which the actual output of the system will belong to each of the groups.

It is important to note that by means of a "source encoder" it is always possible to transform the source output into a sequence of binary digits from which a "source decoder" may generate an acceptable reproduction of

the source output. The importance of this fact rests both upon the simplicity of instrumentation and processing associated with binary representations, as well as the fact that a binary representation (or coding) affords the best characteristics when considering reception under conditions of low signal-to-noise ratio. Should conditions be more favorable (permitting the receiver to distinguish among more than two symbols), the amount of information that can be transmitted per symbol increases considerably by using a higher-radix alphabet. For a new p'ary base, the number of possible N-words (code sequences consisting of N symbols increases from 2^N to p^N . However, the high signal-to-noise ratio (increased symbol distinguishing capability) requirement suggests that those codes are not as versatile in application as the binary ones. Non-binary codes are thus employed whenever it is desirable to trade detection sensitivity for information capacity or security of a code. Because binary codes usually provide adequate capacity and security in practice, non-binary codes have not experienced too much popularity. A separate encoder-decoder pair may also be associated with the channel, and the input of the channel encoder and the output of the channel decoder may, without any loss of generality, be assumed to consist of binary digits regardless of the characteristics of the source and of the user. This implies that the encoding and decoding strategies pertinent to the channel may be planned independently of those pertinent to the source.

7.1 The Binary Detection Problem

In its most elementary form, the binary detection problem may be formulated as follows:

We assume that the time scale is broken up into intervals of duration T seconds, and that the receiver is synchronized to the incoming signal. In any interval $0 < t < T$ the incoming signal is ideally one of two possible narrow-band signals that are identically zero outside the interval. Each signal may be conveniently represented as the product of a complex low-pass modulation waveform and a cisoidal carrier:

$$s_1(t) = x_1(t)e^{j\omega_0 t}, \quad s_2(t) = x_2(t)e^{j\omega_0 t} \quad (15)$$

where ω_0 is a suitably defined carrier frequency. In the simplest form of the problem, the received waveform, $f(t)$, in any interval $0 < t < T$ is assumed to consist of one or the other of the possible signals in (15), together with a stationary, additive, white, gaussian noise of zero mean. Thus

$$f(t) = x_i(t) \cdot e^{j(\omega_0 t + \theta)} + n(t), \quad i = 1 \text{ or } 2, \quad 0 \leq t \leq T \quad (16)$$

where θ is a possible uncertainty in the received carrier phase.

The functions $s_1(t)$ and $s_2(t)$ are assumed to be known in all detail to the receiver (except possibly for an uncertainty about the received carrier phase θ), but are otherwise completely arbitrary. It is further assumed that $s_1(t)$ and $s_2(t)$ are a priori* equally-likely in each interval, and that choices are not correlated from interval to interval. Since any reasonably efficient code will employ an equal number of zeros and ones, the assumption of a priori equal likelihood of $s_1(t)$ and $s_2(t)$ is reasonable. But the assumption that choices are uncorrelated from interval to interval implies that there is no information in the received waveform $f(t)$ outside the

* An a priori probability describes our expectations about the possible presence of signals before receiving $f(t)$; an a posteriori probability describes our state of knowledge after receiving $f(t)$.

the interval $0 < t < T$ that is relevant to the decision about the signal in the interval $0 < t < T$. This may not be strictly true in coded communications, where inter-digit dependence is often met; but systems that take advantage of this dependence (e. g., sequential-detection systems) are quite complex in general, and not easily applicable to the command systems under consideration. Inter-digit independence is then a useful working assumption.

The task of the receiver is to decide, at the end of each interval (e. g., at $t = T$ for the interval $0 \leq t \leq T$) which waveform, $s_1(t)$ or $s_2(t)$, was present during that interval. In order for the receiver to minimize the average number of errors that will be made, it can be shown that for each $f(t)$ received in an interval $0 < t < T$, the receiver should compute the a posteriori probability that $s_1(t)$ was transmitted given that $f(t)$ was received and separately compute the a posteriori probability that $s_2(t)$ was transmitted given that $f(t)$ was received. Then the receiver should announce the presence of the signal whose a posteriori probability is larger.

7.2. Coherent Detection of Nonfading Signals

Coherent detection can be used at the receiver when there is complete knowledge of the received carrier phase θ . In this case it can be shown that, for white gaussian noise, the optimum decision rule based on the a posteriori probability computations simplifies to

$$\int_0^T f(t) \cdot s_1^*(t) dt - \int_0^T f(t) \cdot s_2^*(t) dt \begin{matrix} > s_1 \\ < s_2 \end{matrix} 0 \quad (17)$$

where the star denotes complex conjugate, and the symbol $\gtrless 0$ means compare with zero, announcing s_1 if positive, s_2 if negative. Note that we have assumed that $\theta = 0$. If it is not, a small shift in the time reference at the receiver corresponding to the known θ has the effect of reducing θ to zero.

The operation in (17) can be interpreted in two useful ways. Each of the two integrals is seen to represent a correlation operation between the received signal $f(t)$ and one of the two possible transmitted signals. Thus the first interpretation of (17) is : Cross-correlate the received signal $f(t)$ with $s_1^*(t)$ and $s_2^*(t)$ (the complex conjugates of the transmitted signals) separately, then sample the cross-correlator outputs at time $t=T$, and compare the magnitudes of the two samples. The receiver should announce the presence of the signal used in the branch that produces the larger sample. The cross-correlation operation is realized by a multiplier that multiplies $f(t)$ by the complex conjugated replica of the transmitted signal (s_1^* or s_2^*) which is available at the receiver, followed by a low-pass filter that performs the integration. This is called correlation detection.

The second interpretation of (17) is based upon a comparison of each of the integrals in (17) with the well-known superposition integral of linear system analysis. This comparison suggests the recognition of the integrals as the responses to $f(t)$, sampled at $t=T$, of linear filters whose impulse responses are $s_1^*(T-t)$ and $s_2^*(T-t)$. The receiver then compares the

magnitudes of the two samples, and announces the presence of the signal corresponding to the branch that produces the larger sample. This is matched-filter detection.

The performance of the optimum detection receiver is best characterized by the probability that the decision based upon the above rule is in error. The error probability is strongly dependent on two parameters of the transmitted signals s_1 and s_2 . These are the energy content E of s_1 or s_2 (assumed to be the same) and their correlation coefficient λ , defined by

$$\begin{aligned} E &= \frac{1}{2} \int_0^T |s_i(t)|^2 dt \\ \lambda &= \frac{1}{2E} \int_0^T s_1^*(t) s_2(t) dt \end{aligned} \quad (18)$$

The factor $1/2$ appearing in these definitions ensures that E is the energy in the actual signals, which are the real parts of s_1 and s_2 given in (15). The error probability P_e is found to be

$$P_e = \int_Q^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \quad (19)$$

in which

$$Q = \sqrt{\frac{E}{2N_0}} (1-\lambda)$$

and N_0 is the noise power density. Expression (19) is the familiar error integral. It is seen to depend only on λ and the ratio E/N_0 . Thus for a

fixed λ , P_e is completely determined by the ratio of signal energy E to noise power density N_0 . This is a very interesting result that holds for a wide class of digital communication situations, and implies that the fundamental limitation on the performance of the system is imposed by the noise power density (rather than some related total noise power).

It can be seen from (19) that the lowest possible P_e for a fixed E/N_0 is obtained when $\lambda = -1$. This condition is satisfied only if $s_1(t) \equiv -s_2(t)$, in which case the system is said to employ Phase-Reversal Signaling. On the other hand, signals that satisfy the condition $\lambda = 0$ are called Orthogonal Signals. Notice that in order to achieve a specified P_e in the presence of a given noise density N_0 the signal energy E required for phase-reversal signaling is half that required for orthogonal signaling. This is the often quoted "3 db advantage" of phase-reversal signaling.

It is important to keep in mind the assumptions that must hold in order that the preceding analysis and conclusions (in particular, the one relating to the 3 db advantage of phase-reversal signaling over orthogonal signaling) to hold. Of these assumptions, three deserve careful attention, namely:

- (a) that the binary waveforms are not disturbed by any convolutional perturbations in the channel,
- (b) that the only disturbance present is additive white gaussian noise, and

- (c) that a perfect phase reference is available at the receiver for use in the detection of the PSK signal.

The effect of unavoidable deviations from these assumptions upon achievable performance will be considered in the following sections. In this discussion we must necessarily take into account the basic command-destruct (Range Safety) channel characteristics that will influence performance. We start with the situation in which no reliable knowledge of phase exists that could be used in the detection process.

7.3 Noncoherent Detection of Nonfading Binary Signals

When the receiver has no knowledge of the received carrier phase θ , or when θ undergoes random fluctuations having a uniform distribution $p(\theta) = 1/2\pi$, $< \theta < 2\pi$, it can be shown that the optimum decision rule becomes:

$$\left| \int_0^T f(t) \cdot s_1^*(t) dt \right| \underset{s_2}{\overset{s_1}{>}} \left| \int_0^T f(t) \cdot s_2^*(t) dt \right| \quad (20)$$

Comparison of (20) with (17) shows that the only change introduced into the decision rule by the unknown θ is the operation of taking the magnitude of the integrals. This can be interpreted as an envelope-detection operation preceding the sampler and following the matched filters or cross-correlators. Thus the optimum noncoherent detector consists of the same blocks as a coherent detector built for an arbitrary but fixed θ (say $\theta = 0$) plus the necessary envelope detectors.

The effect of including the envelope-detection operation is to destroy the uncertain phase "information" in the matched-filter outputs, while retaining only the meaningful amplitude information. On this basis, it becomes evident that phase-reversal signaling is entirely unsuitable for use with noncoherent detection, because the phase reversal events would be completely lost in the envelope detectors. For other types of signals, the probability of error with a noncoherent detector turns out to be slightly higher than the probability of error with a coherent detector.

It is interesting to note that the detection schemes outlined above, whether coherent or noncoherent, are not grossly sensitive to deviations from the assumptions made in deriving them. Thus the assumption of whiteness of the additive noise need not be strictly satisfied, provided that the noise power spectral density is reasonably flat over the narrow bandwidth occupied by the signals. Similarly, it has been found both analytically and experimentally that if the matched filters deviated from their optimum forms within wide ranges, the deterioration of the error probability was quite tolerable, corresponding only to a few db reduction in effective transmitted power. This facilitates good engineering compromises between complexity, cost and performance. Notice also that the detection schemes do not require any prior knowledge of any signal or noise levels. Only during the process of establishing initial synchronization is it necessary that the signal and noise levels be carefully

controlled. This is because the decision about whether or not satisfactory synchronization has been established is based upon indications that are dependent upon signal and noise levels. But once synchronization has been established, the system performance becomes dependent upon a comparison of two things that are equally dependent upon the same input levels. Thus, if the signal power drops at the receiver input, the optimum decision rule remains the same, because the comparison operation eliminates the dependence on amplitude. There will be, of course, a drop in E/N_0 if this happens, with a corresponding rise in P_e .

7.4 Binary Detection in the Command-Destruct Receiver Environment

The application of the preceding results to the binary detection problem in the command receiver environment must take into account the vital electromagnetic characteristics of this environment.

First, it is necessary to keep in mind that the results apply as long as no irreversible (or information destroying) operations precede the matched-filter or cross-correlation operations. All nonlinear demodulation processes destroy information below their thresholds of linear output vs. input S/N characteristics. Consequently, irretrievable damage will be incurred if a demodulator is applied below its threshold to the recovery of the noisy binary signals on subcarriers prior to the matched-filter operations.

Second, it is necessary to employ automatic phase or frequency control at the receiver in order to counterbalance the effects of vehicle motion relative to the ground control station. This tracking operation may be secured in a number of ways. One method is to extract from the received signal a pure carrier term containing the doppler-and phase-shift information, and to use this information to make the necessary tracking corrections. Thus, if the carrier term is fed into the input of a phase-locked loop that controls one of the local oscillators in the receiver, it is possible to produce an i-f signal with substantially fixed frequency and phase. Another method is to derive the frequency and phase shift information from a pilot tone transmitted in conjunction with the command code information (either by using one of the command channels exclusively for the purpose, or by introducing the pilot tone in some appropriate location in the transmitted r-f or baseband spectrum) and then use that information to make the corrections. Each of these techniques is effective as long as any fading encountered in the channel is flat, i. e., uniform over all frequencies of importance in the structure of the signal, and the fading mechanism does not destroy the phase coherence among the signal components. As we have already noted, attenuation caused by the plasma in the booster exhaust, though it may be flat over all frequencies of interest, will be accompanied by destruction of phase coherence among the signal components.

It is important to observe that with noncoherent detection (which is not applicable to phase-reversal signaling) the above phase correction is not needed as long as the signal is maintained within the receiver pass-band by coarse AFC action.

The presence of "plasma fading" and/or of changing-multipath interference fading of any kind will also have detrimental effects upon binary detection. The presence of even a moderate amount of these "convolutional" disturbances will cause a marked deterioration in the performance of any binary detection system; but orthogonal ($\lambda = 0$) binary signals result in much better performance than phase-reversal signals and normally little is gained by the use of a coherent receiver unless a very elaborate system of phase stabilization is employed. This conclusion holds whether the binary waveforms modulate the r-f carrier directly or are conveyed in some sub-carrier or baseband modulation scheme.

Some techniques have been proposed to save the performance of phase-reversal signaling using coherent detection from sharp degradation by multipath and other fading. All of these techniques strive to supply the receiver with a sufficiently stable phase reference in the presence of the convolutional perturbations to alleviate the major degradation otherwise introduced by these disturbances. These techniques, however, are limited in their effectiveness by the speed with which the phase correction can be accomplished at the receiver, and thus by the fastest allowable fading.

The simplest of these techniques is to equip the receiver with sufficiently fast phase-tracking circuitry to enable it to correct for the random phase fluctuations due to fading and other interference, and thus present a signal with a stable carrier phase to the coherent detectors. Obviously, such a scheme cannot compensate for rapid fading that approaches the command code bit rate in speed.

Another technique is to use "incomplete" phase-reversal signaling, i. e., a little less than complete negative correlation. This amounts to transmitting a small unmodulated subcarrier component all the time, and using it at the receiver to make instantaneous corrections of the received subcarrier phase. The main disadvantage of this approach relative to the complete phase reversal technique is that the transmitted power lost in the unmodulated component offsets most or all of the 3 db advantage brought out in Section 7.2 above.

The technique that has been used successfully in other applications is called differentially-coherent detection. It consists essentially of making a comparison of the phase of the carrier or subcarrier in each detection interval with the phase in the immediately preceding interval. A detected change in phase would then be interpreted as a change from a mark to a space (or vice versa), while no change in phase means a succession of identical bits. It is clear that, if the fading or the interference is sufficiently rapid so that the carrier or subcarrier phase can change appreciably between successive detection intervals, the performance of this scheme will

degenerate to that of a coherent detector with a noisy phase reference. Once again, the fading must be much slower than the bit rate of the command code for this scheme to show improvement.

Finally, we can conclude that the best choice for a secure command system using binary data transmission which is to operate under both fading and nonfading conditions is to use orthogonal signals and non-coherent detection. The other signaling and detection schemes that have been suggested as better alternatives are limited in their applicability to fading rates that remain much slower than the bit rate. This does not exclude the possibility of the future development of more sophisticated phase stabilization techniques that could perform effectively under more rapid fading and interference conditions. Since the plasma attenuation and phase randomization and the changing-multipath fading are largely peculiar to near-space missions, phase-reversal signaling should not be discounted entirely for medium and deep space missions unless careful study of the medium and deep space channels reveal intolerable phase instabilities or noticeable reductions in coherence bandwidth, and unless the presence of other multiplexed command signals in the baseband causes spurious components to be superimposed upon the desired code spectrum. Any desire for standardization would, however, favor orthogonal signaling with incoherent detection, especially when a wide variety of channel and electromagnetic environment conditions is expected, and when other command signals are expected to be simultaneously present in the same channel or the same baseband.